HighSpeed TCP for Large Congestion Windows

Evandro de Souza

TCP Congestion Control

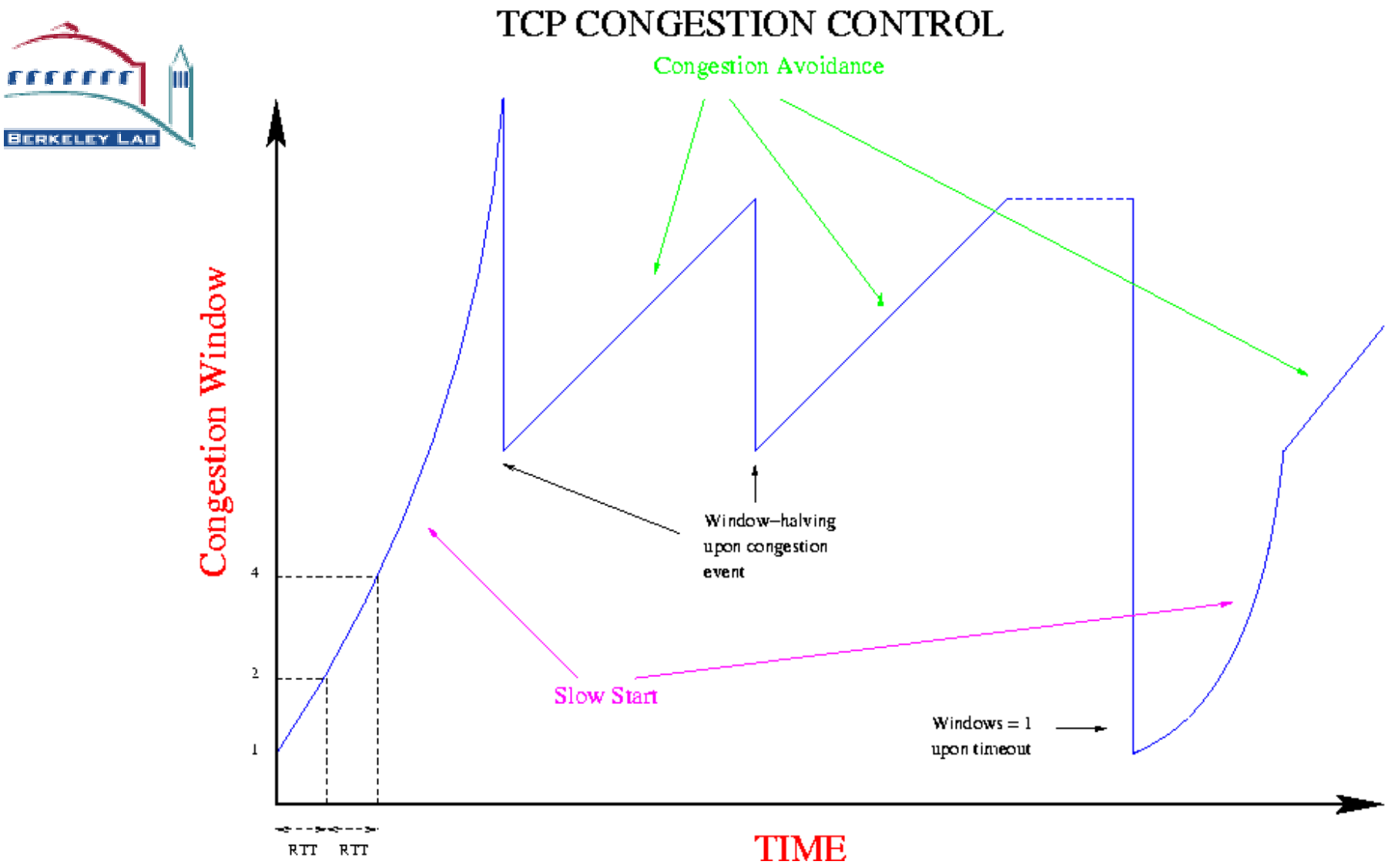
It is a mechanism that controls TCP sending rate. It adopts a window-based flow control, which controls the number of on-the-fly packets in the network.

The window size is updated at the receipt of ACK (ACKnowledgement) packet

TCP Congestion Control

The source terminal is allowed to send the number of packets given by its window size (cwnd)

The key idea of the congestion control mechanism of TCP is to dynamically control the window size according to severity of the congestion in the network


TCP CONGESTION CONTROL
Congestion Avoidance
Congestion Window
Window–halving
upon congestion
event
Slow Start
Windows = 1
upon timeout
4
2
1
RTT
RTT
TIME

TCP Response Function

It has been known that the current TCP does not scale to high-speed networks directly

TCP's response function:

$$S = \frac{\sqrt{1.5}}{\sqrt{p}} \; pkts \; / \; RTT$$

High Speed Problem

For example, for a TCP connection with 1500-byte packets and a 100 ms round-trip time, filling a 10 Gbps pipe would require a congestion window of 83,333 packets, and a packet drop rate of at most one drop every 5,000,000,000 packets. This is at most one drop per S = 6000s, or 1h:40m

HSTCP Proposal

A modification to TCP's congestion control mechanism for use with TCP connections with large congestion windows.

It aims to achieve high per-connection throughput without requiring unrealistically low packet loss rates and TCP-compatible performance in environments with moderate or high congestion

AIMD Algorithm

Basic algorithm to adjust the TCP congestion window (Von Jacobson 88)

- ACK: w <-- w + a/w
- DROP: w <-- w - b*w
- SLOW-START: w <-- w + c
- a = 1; b = 0.5; c = 1

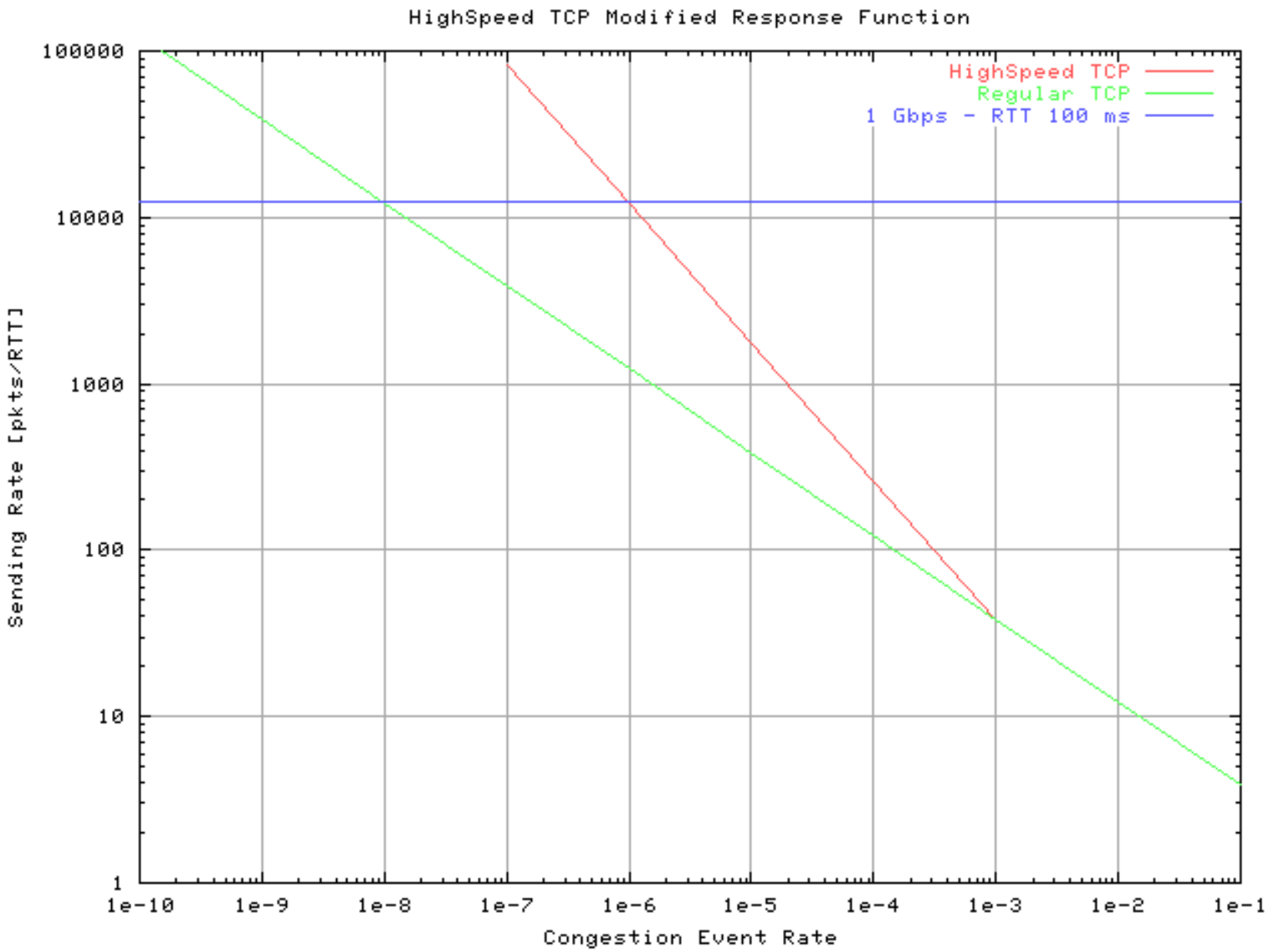
HSTCP Algorithm

HSTCP algorithm:

- ACK: w <-- w + a(w)/w
- DROP: w <-- w - b(w) * w

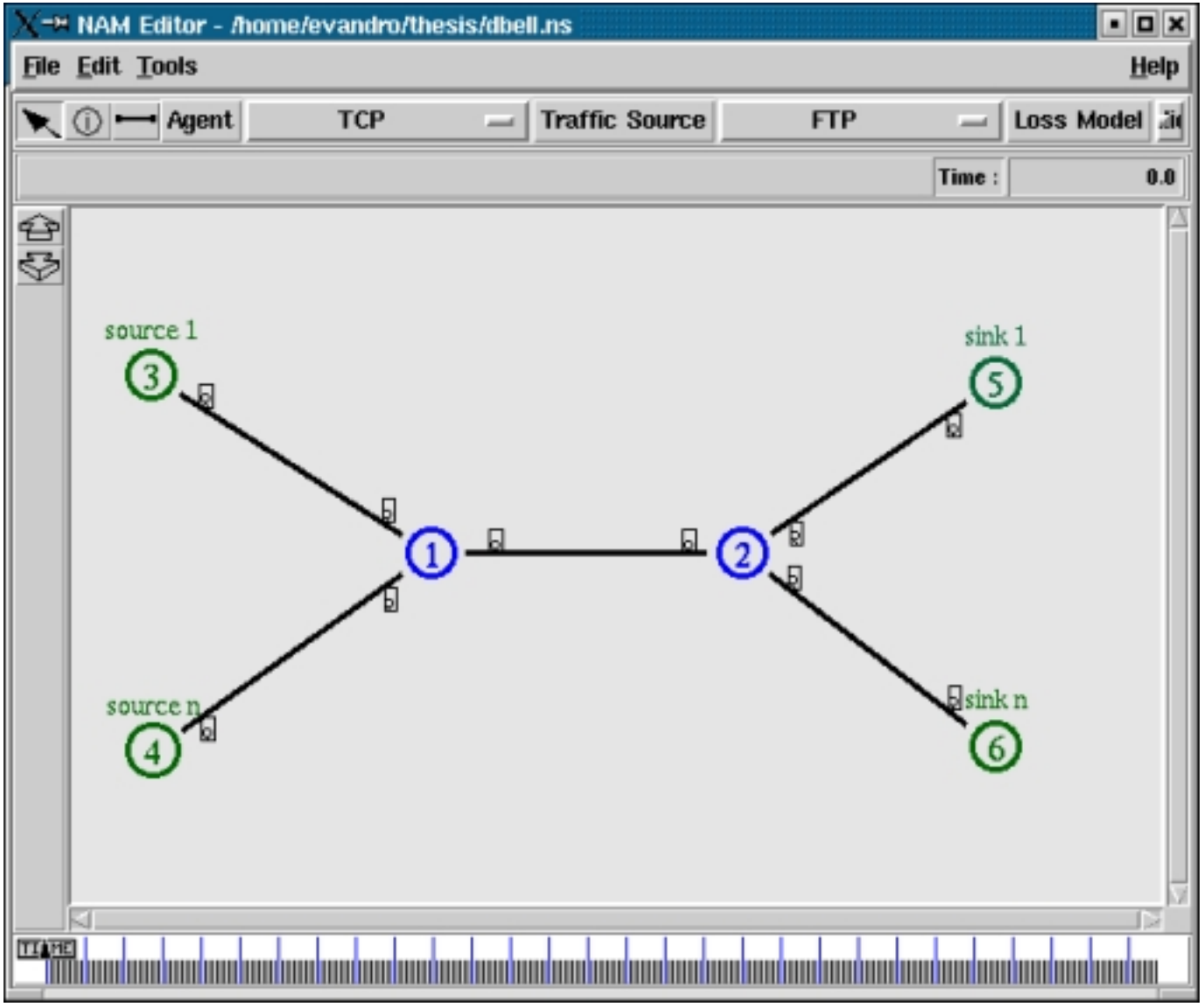
$$a(w) = \frac{HWin^2 \times HP \times 2 \times b(w)}{2 - b(w)}$$

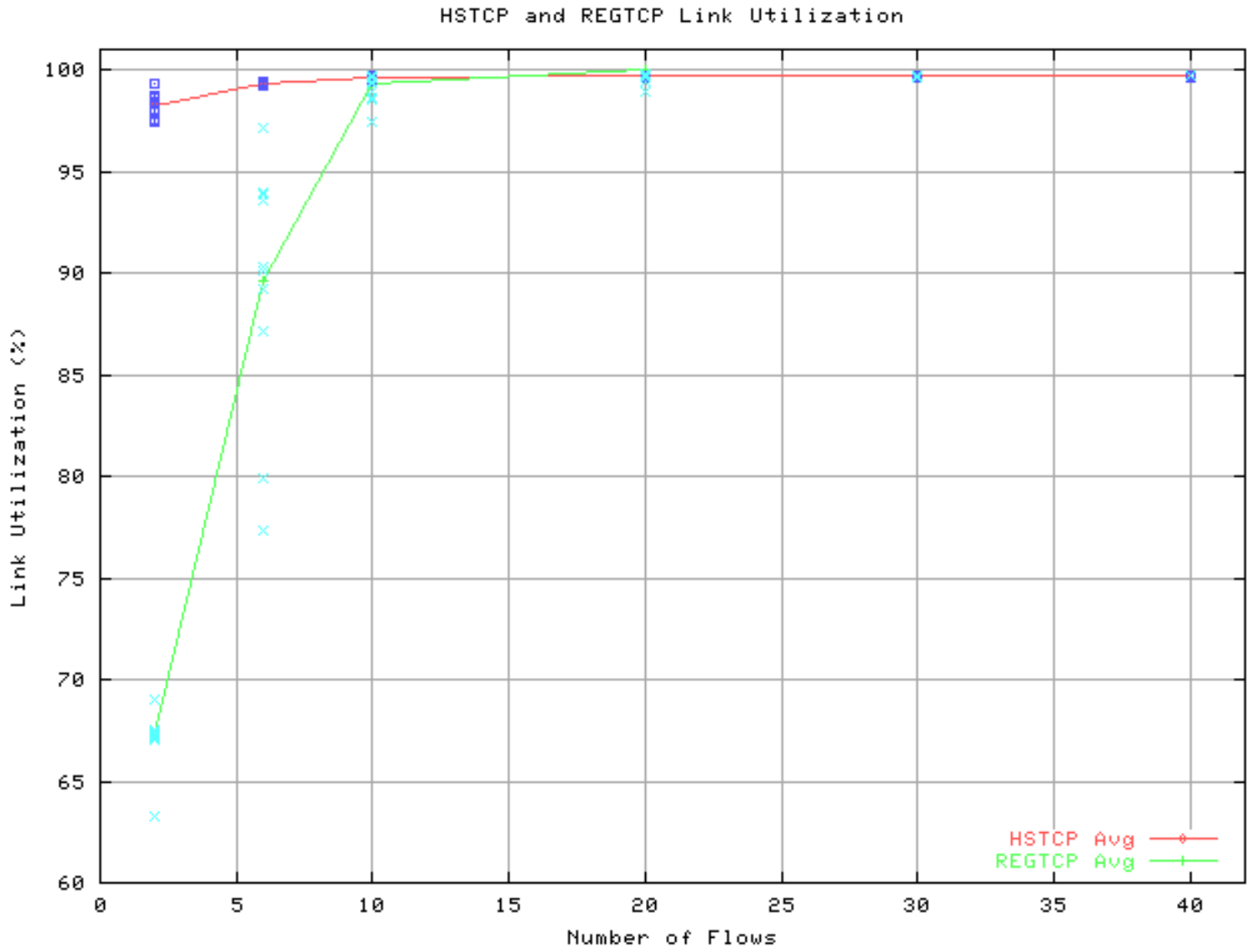
$$b(w) = \frac{(HDecr - 0.5) \times (\log(w) - \log(W))}{\log(W1) - \log(W)} + 0.5$$

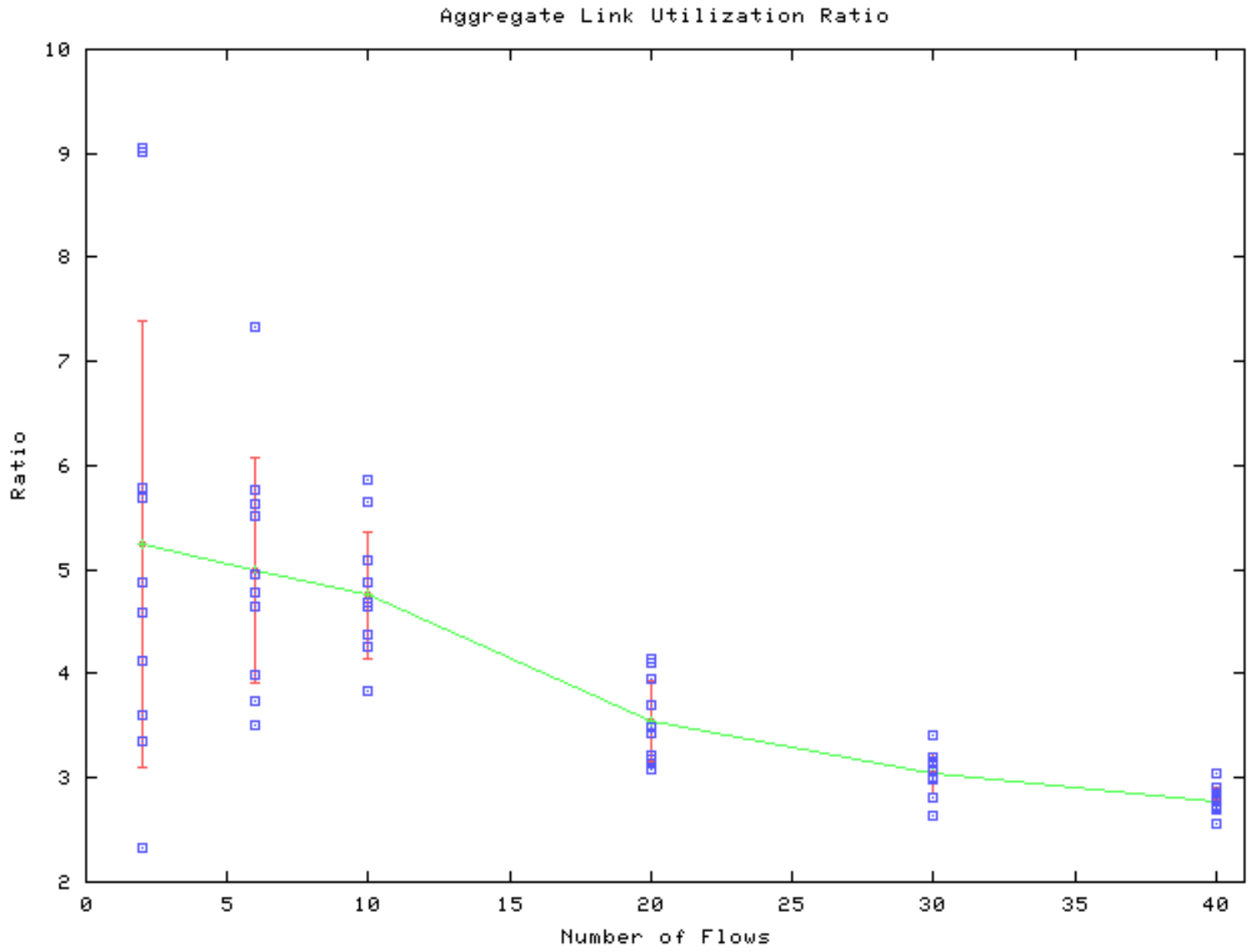

HighSpeed TCP Modified Response Function
HighSpeed TCP
Regular TCP
1 Gbps - RTT 100 ms
Sending Rate [pkts/RTT]
100000
10000
1000
100
10
1
1e-10
1e-9
1e-8
1e-7
1e-6
1e-5
1e-4
1e-3
1e-2
1e-1
Congestion Event Rate

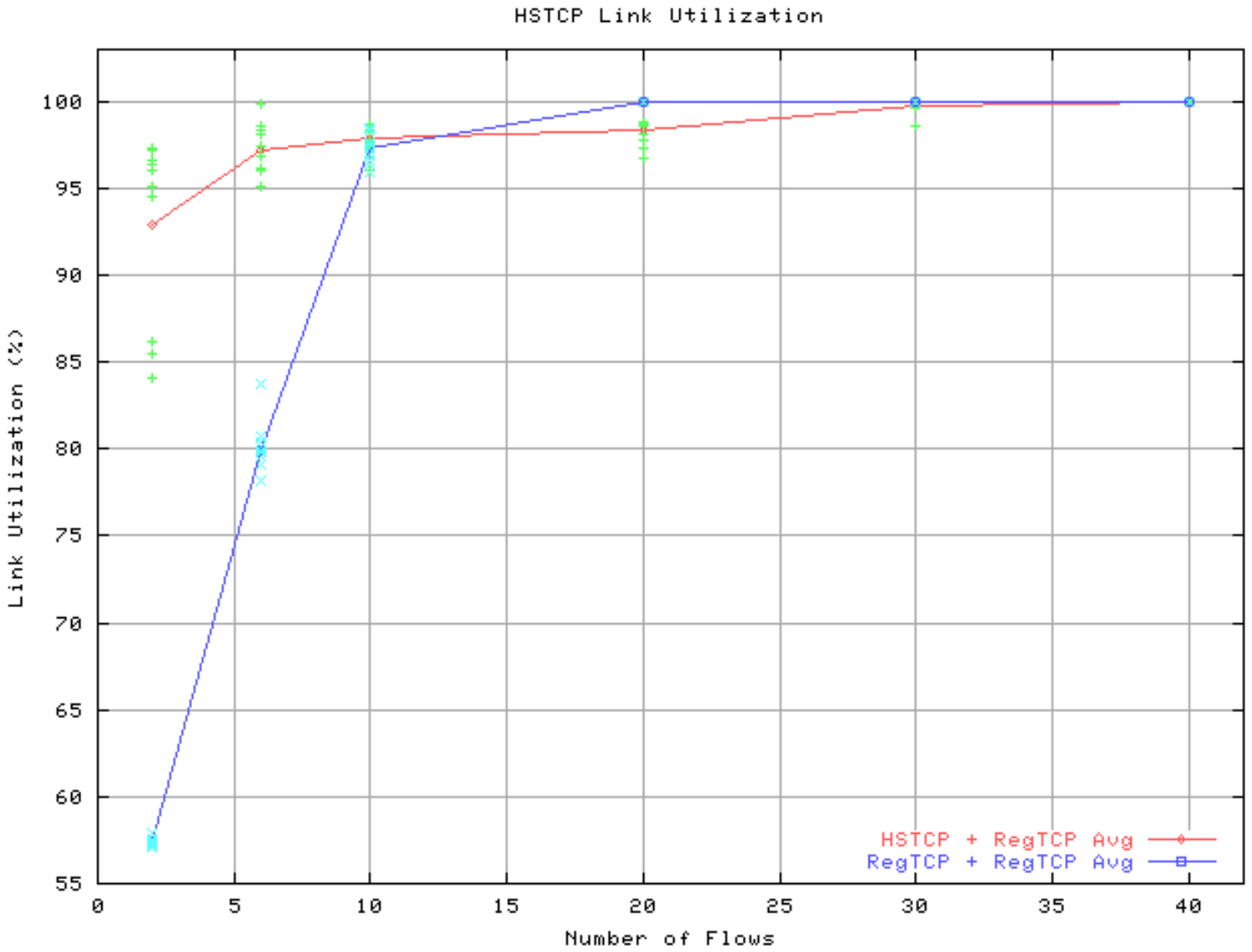
Objective

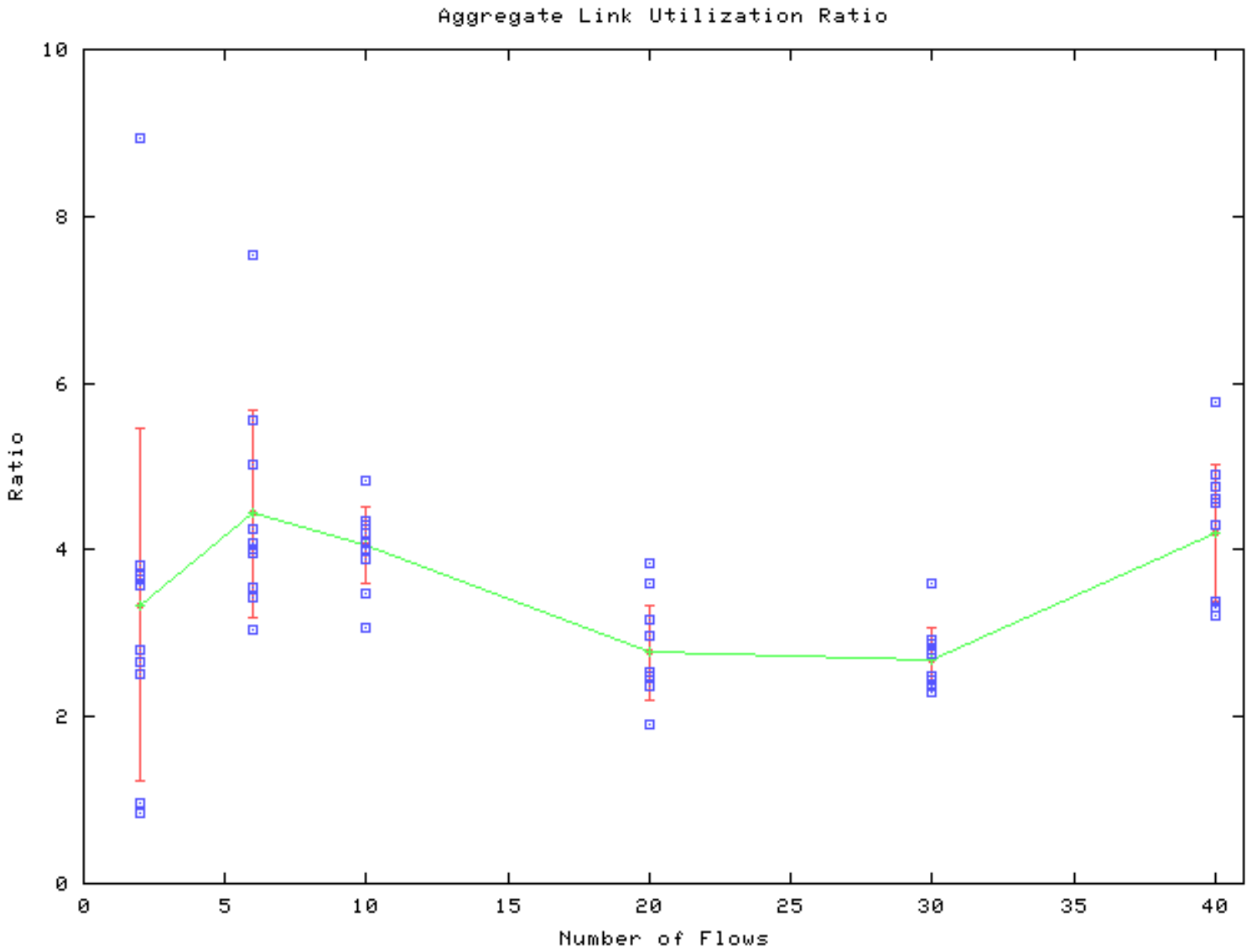
Study the deployment of HSTCP in high speed long distance links as an effective mechanism for bulk data transfer, while maintaining fairness with other types of TCP already in use.

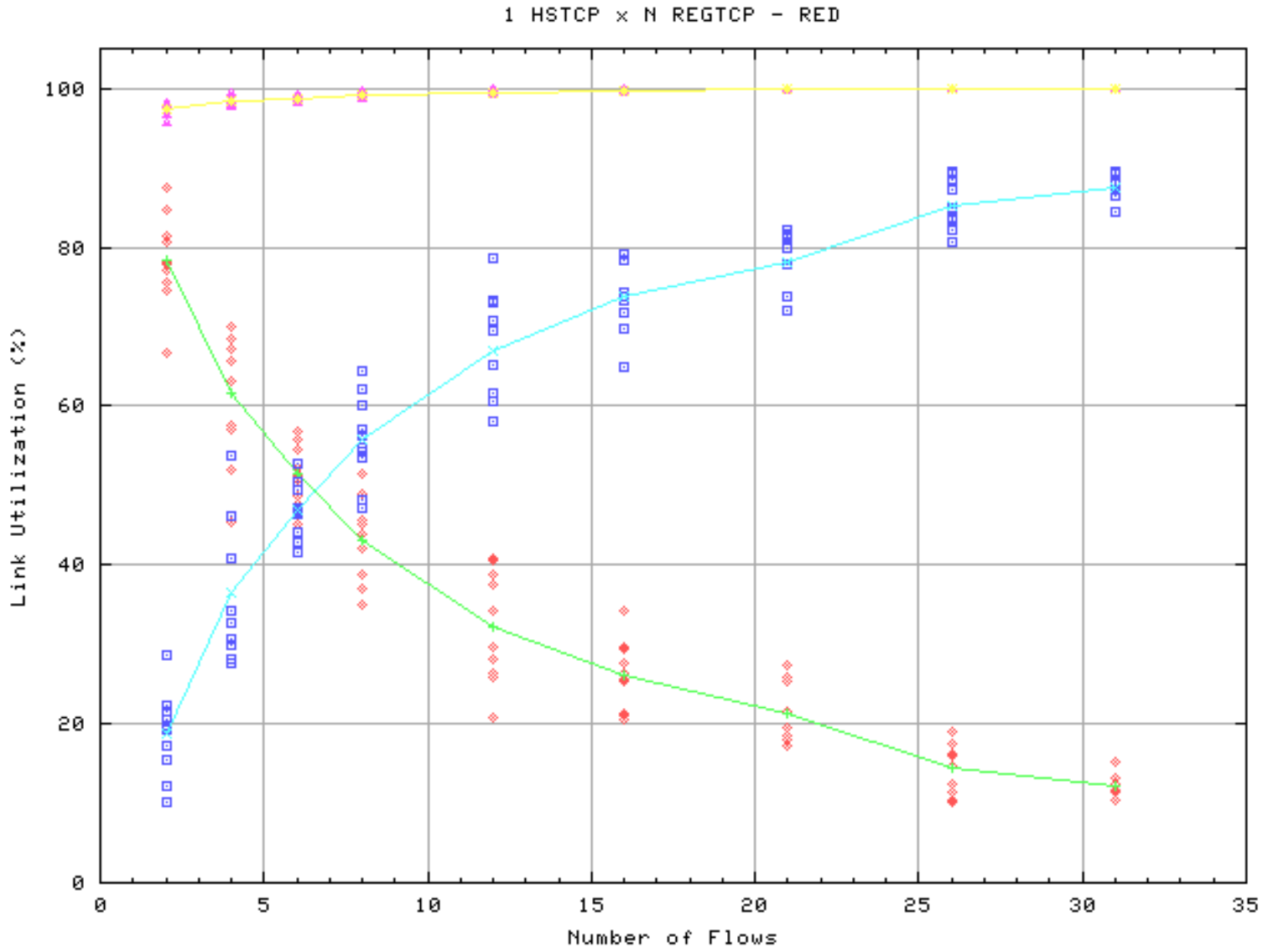

NAM Editor - /home/evandro/thesis/dbell.ns
File Edit Tools
Help
Agent
TCP
Traffic Source
FTP
Loss Model
Time :
0.0
source 1
sink 1
source n
sink n
1
2
3
4
5
6

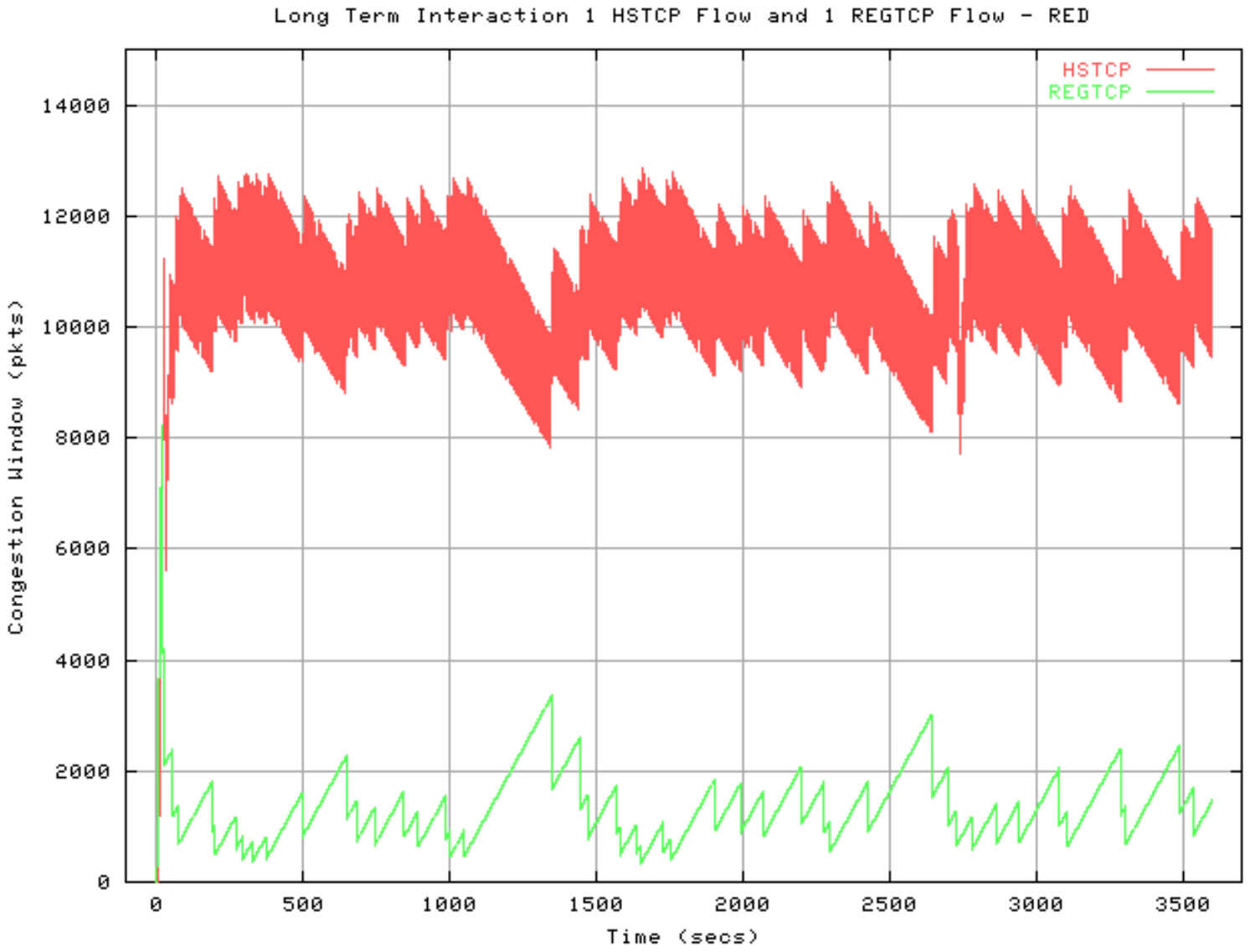

HSTCP and REGTCP Link Utilization
Link Utilization (%)
Number of Flows
100
95
90
85
80
75
70
65
60
0
5
10
15
20
25
30
35
40
HSTCP Avg
REGTCP Avg

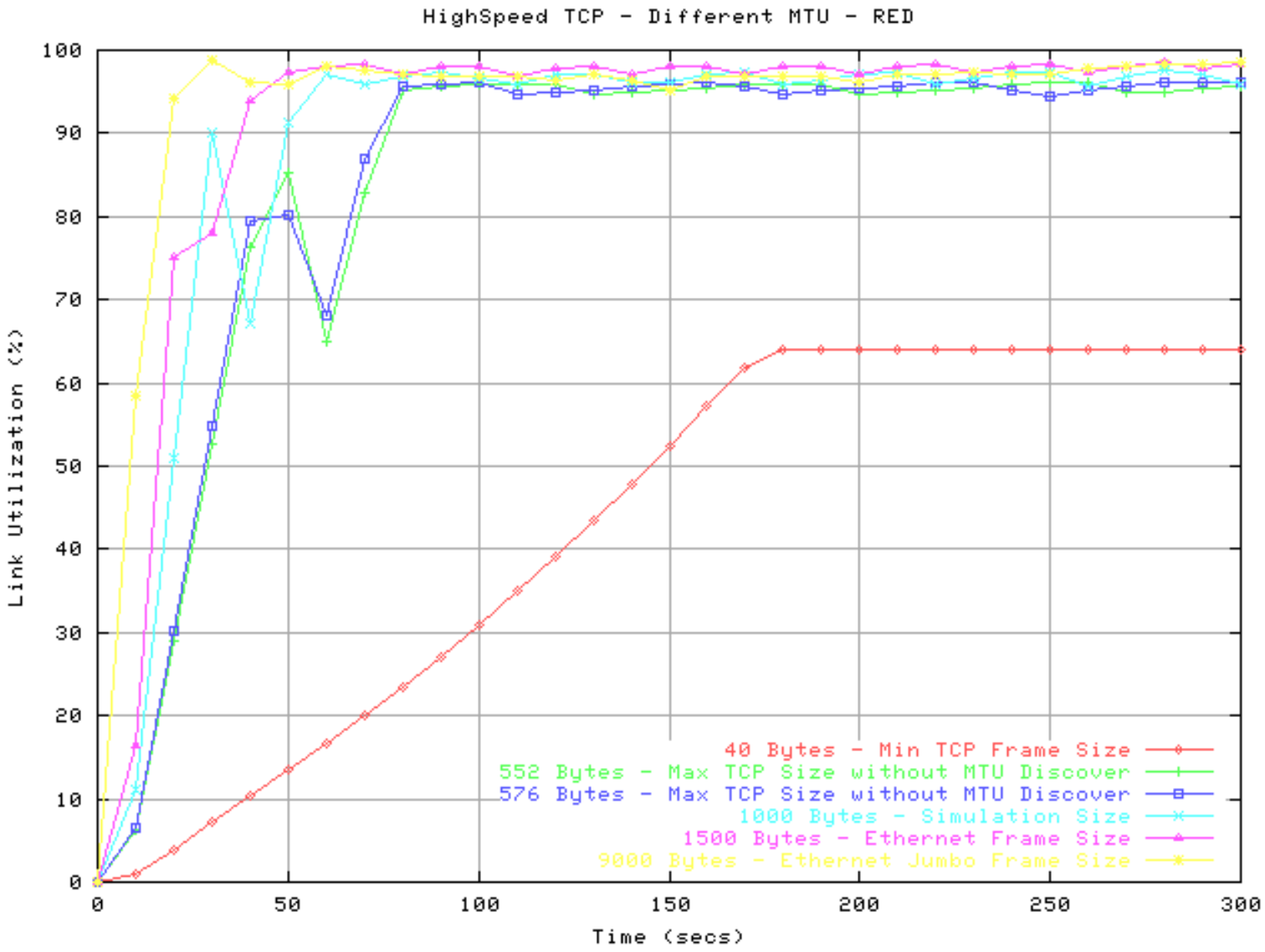

Aggregate Link Utilization Ratio
Ratio
Number of Flows
10
9
8
7
6
5
4
3
2
0
5
10
15
20
25
30
35
40


HSTCP Link Utilization
Link Utilization (%)
Number of Flows
100
95
90
85
80
75
70
65
60
55
0
5
10
15
20
25
30
35
40
HSTCP + RegTCP Avg
RegTCP + RegTCP Avg


Aggregate Link Utilization Ratio
Ratio
Number of Flows
0
2
4
6
8
10
0
5
10
15
20
25
30
35
40


1 HSTCP x N REGTCP - RED
Link Utilization (%)
Number of Flows
0
20
40
60
80
100
0
5
10
15
20
25
30
35


Long Term Interaction 1 HSTCP Flow and 1 REGTCP Flow - RED
HSTCP
REGTCP
Congestion Window (pkts)
14000
12000
10000
8000
6000
4000
2000
0
0
500
1000
1500
2000
2500
3000
3500
Time (secs)


HighSpeed TCP - Different MTU - RED
Link Utilization (%)
Time (secs)
0
10
20
30
40
50
60
70
80
90
100
0
50
100
150
200
250
300
40 Bytes - Min TCP Frame Size
552 Bytes - Max TCP Size without MTU Discover
576 Bytes - Max TCP Size without MTU Discover
1000 Bytes - Simulation Size
1500 Bytes - Ethernet Frame Size
9000 Bytes - Ethernet Jumbo Frame Size

Present Investigations

Behavior with bursty traffic

Comparison with multiple TCP streams

Effects of competing in LAN